CF "Common Concept":
What? Why? How?
and What Next?

Alison Pamment   Alison.Pamment@stfc.ac.uk

British Atmospheric Data Centre

NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL

CF Common Concept: What? (1)

In a nutshell:

- A proposal to extend the CF conventions to allow metadata bundles to be defined and labelled

- Each bundle must contain at least a standard_name and (optionally) other attributes that help to identify the physical quantity represented by the data, e.g., cell_methods, scalar coordinate variable, coordinate bounds

CF Common Concept: What? (2)

- For example, a 2m temperature would be described by a number of CF attributes:

```
variables:
double height
height: standard_name = "height" ;
height:units = "m" ;

float  t2m(ntimes) ;
    t2m:standard_name = "air_temperature" ;
    t2m:units = "K" ;
    t2m:coordinates = "height" ;

    data:
    height  2.0 ;
```

CF Common Concept: What? (3)

- The "concept" is the definition of the attribute bundle, e.g., which CF attributes and values adequately describe 2 metre temperature

- It is "common" because the same concept may occur in many datasets held at many data centres and in many science domains.

CF Common Concept: What? (4)

- It is proposed that a single common concept definition can be given multiple labels (scoped names) which would exist in namespaces:

  {badc.nerc.ac.uk}2m_air_temperature

  {wdc-climate.de}air_temperature_at_2m

  {gfdl.noaa.gov}airTemp2m

- Each definition will be assigned a URN

British Atmospheric Data Centre

NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL

CF Common Concept: Why? (1)

- Not all scientists want to use the same terminology to describe the same concept, e.g., T2m, 2mtemp, temp2.

- Data centres can serve their user communities using familiar terminology – not everyone speaks CF (or wants to!)

- Map legacy data to CF

- Deals with long-standing need for abbreviations for standard names

CF Common Concept: Why? (2)

Use Cases

- High cloud

- Frost days

- sea_surface_wind_wave_mean_period_from _variance_spectral_density_first_frequency_ moment

- But note that Common Concept is NOT intended as a replacement for standard names

British Atmospheric Data Centre

NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL

CF Common Concept: How? (1)

- An additional CF attribute would be added to the file:

variables:

double height

height: standard_name = "height" ;

height:units = "m" ;

float t2m(ntimes) ;

    t2m:standard_name = "air_temperature" ;

    t2m:units = "K" ;

    t2m:coordinates = "height" ;

    t2m:common_concept =

    "{badc.nerc.ac.uk}2m_air_temperature:urn:cf-cc:blah123" ;
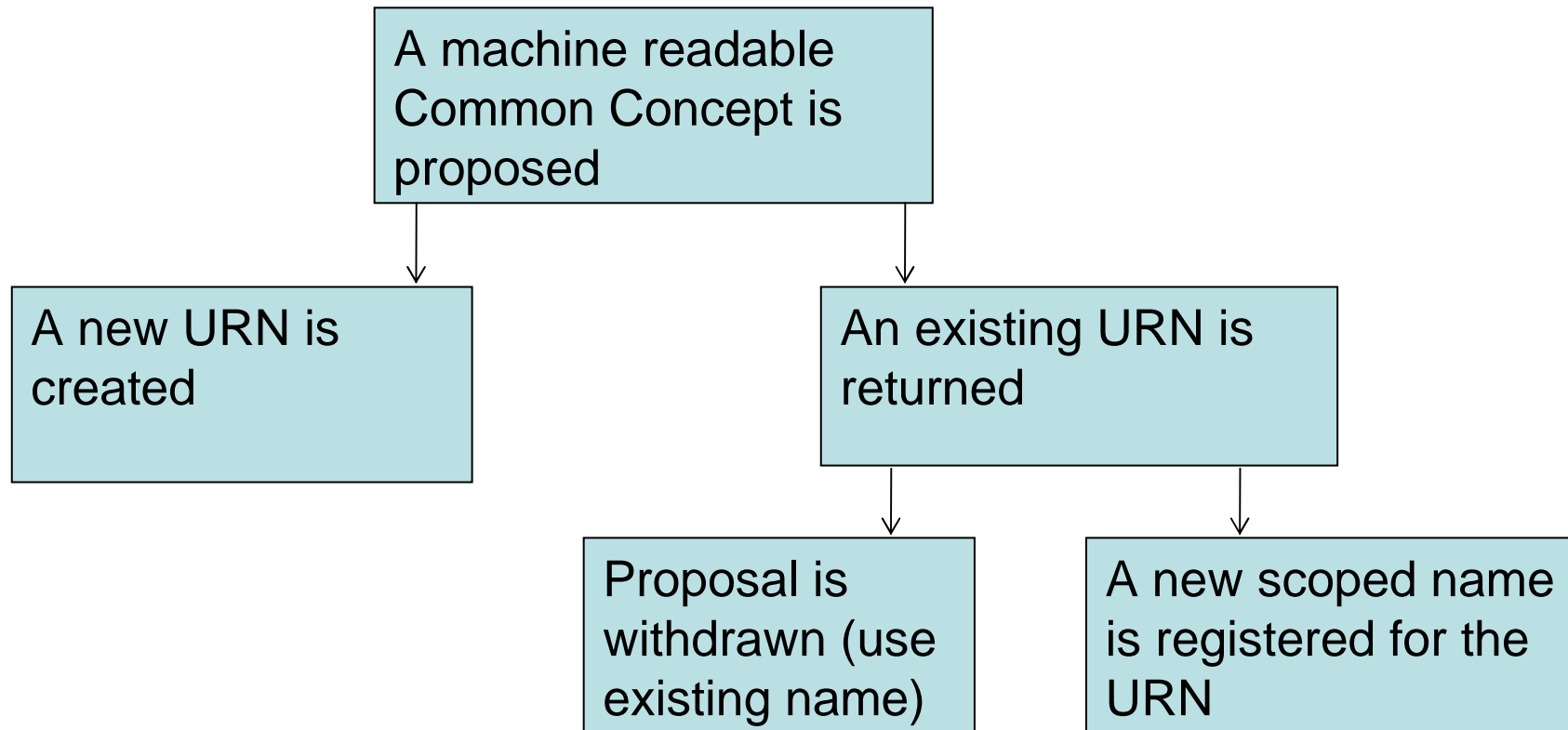
    data:

    height 2.0 ;

CF Common Concept: How? (2)

- Proposals should be made to the CF mailing list

- Proposals for Common Concepts should be machine readable and consist of:

  - A suggested name in the format of {namespace}scopename

  - A list of pseudo-cdl formatted constraints

  - A plain text description of the constraints

CF Common Concept: How? (3)

A machine readable Common Concept is proposed

A new URN is created

An existing URN is returned

Proposal is withdrawn (use existing name)

A new scoped name is registered for the URN

CF Common Concept: How? (4)

- Define a bundle using CDL - CF trac ticket 29

E.g. "high cloud"

```
dimensions:

hgt  =  1  ;

variables:

float  x(unconstrained)  ;

x: standard_name = "cloud_area_fraction"  ;

x: units = "1"  ;

x:coordinates  =  "height + unconstrained"  ;

x:common_concept :"{gfdl.noaa.gov}high_cloud;tbd"  ;

 float height(unconstrained)  ;

height: valid_min = 7000.  ;

height:valid_max  =  14000.  ;

height: units = "m"
```
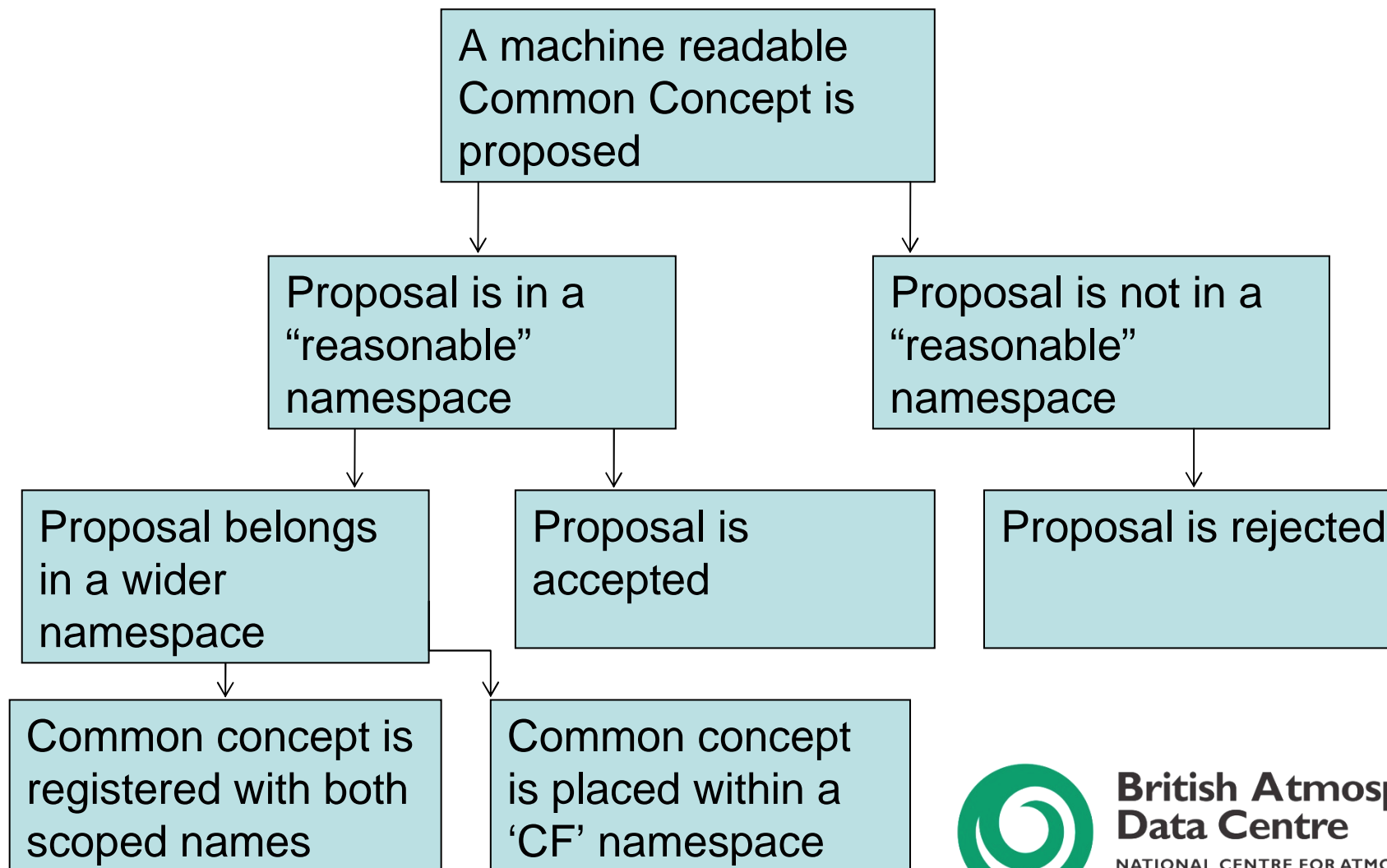
CF Common Concept: How? (5)

CF Common Concept: How? (6)

- What syntax should be used for the URIs?

- General agreement that URIs should be opaque, i.e. contain no additional semantics

- Do we need a CF namespace?

- Should there be a list of permitted external conventions/namespaces?

- Trac ticket 27 addresses the question of identifying namespaces and hence scoped attributes

CF Common Concept: What (again)?

- Should common_concept attribute contain:
  - Scoped name only (no namespace) (2 votes)
  - URI only (1 vote)
  - Both URI and scoped name (2 votes)

- Should it be split into two attributes?

- General agreement on using opaque URIs.

CF Common Concept – What Next?

- Common Concept originated November 2007

- Proposed to CF early 2008

- Moderator's summary August 2008

- Need to produce a second draft of the proposal (it will have some gaps!)

- Ticket 27 (URNs) (Ethan Davis has agreed to moderate)

- Ticket 29 needs a moderator

CF Common Concept

- https://cf-pcmdi.llnl.gov/trac/ticket/24

- https:cf-pcmdi.llnl.gov/trac/ticket/27

- https://cf-pcmdi.llnl.gov.trac/ticket/29